

Sampling-helper: A web-based tool to assess the reliability of sampling strategies in sewers and receiving waters

Sampling-Helper: un outil internet pour qualifier la représentativité de stratégies d'échantillonnage en réseaux d'assainissement et milieux récepteurs

Rossi L, Rumley L, Ort C*, Minkkinen P** , Barry DA, Chèvre N***

Ecole Polytechnique Fédérale de Lausanne (EPFL), Ecological Engineering Laboratory (ECOL), ENAC – IEE, Station 2, 1015 Lausanne

* The University of Queensland, Advanced Water Management Centre (AWMC) Brisbane 4072, Australia

** Lappeenranta University of Technology, P O Box 20, FIN-53851 Lappeenranta, Finland

*** IMG, Faculty of Geosciences and Environment, University of Lausanne, 1015 Lausanne, Switzerland

RÉSUMÉ

L'échantillonnage est une étape clé dans l'analyse de composés physico-chimiques. Cette étape est particulièrement importante dans le domaine environnemental, par exemple pour les eaux usées, les eaux de ruissellement ou les milieux récepteurs dans lesquels les débits et les concentrations varient fortement au cours du temps. Contrairement aux efforts réalisés spécifiquement en chimie analytique, les développements dans le domaine de l'échantillonnage restent rares. Cette étape peut néanmoins générer des erreurs dans un ordre de grandeur bien supérieur à celui analytique. Dans cet article, nous proposons un outil internet, basé sur une théorie de l'échantillonnage pour identifier et quantifier les erreurs dans les différents processus de prise d'échantillons. Cette théorie générale de l'échantillonnage, déjà appliquée à différents domaines, permet notamment de donner des réponses à des questions liées au nombre d'échantillons nécessaires, à leur volume, à leur représentativité, etc. L'utilisation d'internet pour héberger cette application est destinée à faciliter l'utilisation des outils théoriques et de sensibiliser les professionnels aux incertitudes liées à l'échantillonnage. Un exemple est présenté, qui met en évidence l'importance de l'étape d'échantillonnage dans la qualité des résultats d'analyse.

MOTS CLÉS

Echantillonnage, représentativité, erreurs, incertitudes, hétérogénéité

ABSTRACT

Sampling is a key step in the analysis of chemical compounds. It is particularly important in the environmental field, for example for wastewater effluents, wet-weather discharges or streams in which the flows and concentrations vary greatly over time. In contrast to the improvements that have occurred in analytical measurement, developments in the field of sampling are less active. However, sampling errors may exceed by an order of magnitude those related to analytical processes. We proposed an Internet-based application based on a sampling theory to identify and quantify the errors in the process of taking samples. This general theory of sampling, already applied to different areas, helps to answer questions related to the number of samples, their volume, their representativeness, etc. The use of the internet to host this application facilitates use of theoretical tools and raise awareness of the uncertainties related to sampling. An example is presented, which highlights the importance of the sampling step in the quality of analytical results.

KEYWORDS

Sampling, reliability, errors, uncertainties, heterogeneity

1 INTRODUCTION

Chemical analyses are carried out for purposes such as process control, product quality control and environmental control. The reliability of information is of particular importance in environmental applications to assess risk or to understand contamination processes. Moreover, in the development of any modelling approach it is important that the model output is calibrated/validated against observed data. Uncertainties in observed data generate uncertainties in model predictions. Implicitly (rarely explicitly), the term “analysis” is usually considered as a whole, including the sampling and analytical measurements. Rarely described is how samples were taken. Reported measurement uncertainty often accounts only for that stemming from the chemical analyses (Ort and Gujer, 2006). In recent years much effort has focused on the analytical part of environmental sample analysis, especially for emerging pollutants like pharmaceuticals or biocides (Chèvre et al., 2004) with much equipment investment. However, corresponding investment in sampling equipment are negligible by comparison. The importance of sampling in terms of errors has been demonstrated in numerous domains, and also in environmental areas (Christakos and Olea, 1992; Littlewood, 1995; Isidoro et al. 2003; Bertrand-Krajewski et al. 2003; Johnes, 2007; Ort and Gujer 2006). For example, large uncertainties (30%-100%) in stormwater data result from the different sampling and analytical steps (sampling of water, sample conservation, reconstitution of a flow-weighted composite sample and chemical laboratory analyses) (Rossi 1998). Depending on the sampling strategies, global errors could be even bigger. For example, measurement of trace pollutants in a sewer system may generate error in the range -97 to 166%, without considering other sources of uncertainties than sampling intervals (Ort and Gujer, 2006). The errors linked with sampling steps are difficult to consider due to the lack of knowledge in both the processes of error generation and the definition of these errors. In the environmental domain, water quality monitoring programs are conducted worldwide on rivers, streams and wastewater treatment plants (WWTP) by local authorities. These monitoring programs need to be carefully designed to take account of evidence of fluctuations in the rate and speciation of contaminant transport. There is an urgent need to develop practical tools to improve the reliability of these sampling plans.

Sampling has been considered in the mineral processing domain, where sampling errors may have large economic consequences (Gy, 1988) and a theory of sampling (TOS) has been developed by Pierre Gy. This approach has been shown to be applicable to other domains (e.g., Minkinen, 2004; Paakkunainen et al. 2007). The TOS encompasses all facets of sampling: how errors are generated, how they can be eliminated or reduced and how the residual error can be estimated. However, the theory includes perhaps non-familiar concepts like heterogeneity and non-trivial mathematical developments. The goal of this paper is to present a web-application of the TOS (www.samplinghelper.com) that allows for optimization of sampling and the analytical measurement protocol. The web-application provides a guided way to estimate errors linked with sampling processes for environmental purposes. Although the focus is on water samples, the application domain is broad. Careful optimization of the sampling and measurement steps of the complete analytical procedure may result in considerable savings or in improvement of the reliability of results. An illustration of TOS will be presented for a combined sewer system.

2 METHODOLOGY

The TOS elaborated by Gy (1988, 1992, 1996) and Pitard (1993, 2009) has adapted to water sampling. One contribution of this theory is the mathematical definition of heterogeneity. At different scales, heterogeneity is always present in the environment: It is time-dependent and can be considered as the major cause of sampling errors.

The *Samplinghelper* website permits estimation of the global estimation error (*GEE*), as illustrated in Figure 1, which illustrates the sources of error of an analytical determination according to the TOS. For convenience, the original terminology has been adapted to the environmental domain. The global estimation error is the sum of the total sampling error (*TSE*), the point materialization error (*PME*) and the total analytical error (*TAE*). The user is guided through the different steps illustrated in Figure 1 to estimate the *GEE*. Six web pages allow estimation of the possible errors. As a prerequisite, information on typical pollutant dynamics and flow rate is necessary. The goal here is to estimate concentration and flow changes over time. Indeed, as cited in ISO 5667-1 (1980): “The times and frequencies of sampling ... can be properly decided only after detailed preliminary work, in which a high sampling frequency is necessary to provide the information to which statistical techniques may be applied”. This information can be extracted from previous studies, literature or modelling results.

Different scenarios are already implemented in a database to facilitate use of the application. Otherwise, an import function is provided so that the user can submit a data file (time, flow rates and concentrations).

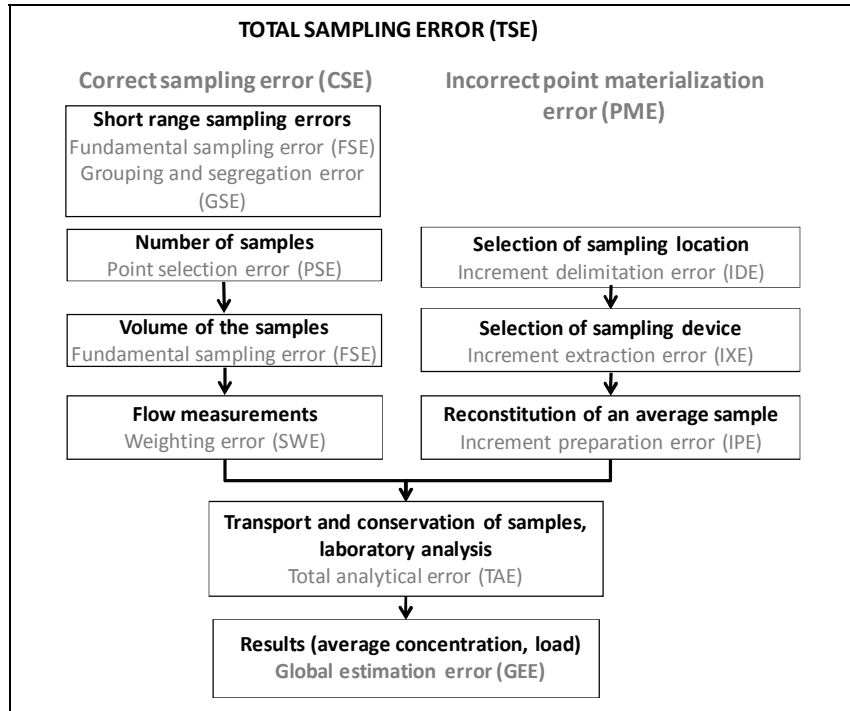


Figure 1: Illustration of the flow chart used in the Web application to generate the global estimation error (*GEE*). The text in greyscale refers to the original terminology of the TOS, that in black to a version for environmental applications.

2.1 Definition of the different errors

Definitions of the different possible errors refer to the problem of contamination of both urban and receiving waters, as detailed in the TOS (Gy, 1996, Pitard, 1993).

2.1.1 Point selection error (PSE)

In rivers and sewers, concentrations and flows fluctuate with time. The number of samples required to characterize the pollutant mass and fluxes increases when the concentrations and/or flows vary greatly. Sampling frequency is calculated using variographic analysis. Specifically, variographic analysis is used to characterize the standard deviations of the sampling errors with different sampling intervals. It accounts for time series auto-correlation. In our application, the variographic analysis produces a graph that illustrates the sampling selection error (standard deviation) as it varies with sample number. Three different sampling modes are considered: systematic sampling (e.g., hourly), stratified sampling (one random sample daily) or “random samples” that can refer to grab samples taken monthly, for example. Note that stratified sampling is rarely used in water pollution studies. If there is no auto-correlation, all three sampling schemes give the same sampling variance $s^2(PSE)$. For systematic sampling, the *PSE* is calculated with:

$$s^2(PSE)_{SY} = \frac{W(j)_{SY}}{n} \quad [1]$$

Where n is the number of samples and $W(j)_{SY}$ is the error-generating function for systematic sampling defined as:

$$W(j)_{SY} = 2w\left(\frac{j}{2}\right) - w'(j) \quad [2]$$

Here, $w(j)$ and $w'(j)$ represent the first order and second order average integrals of the variogram,

defined as:

$$w(j) = \frac{1}{j} \int_0^j V_h(j') dj' \quad [3]$$

$$w'(j) = \frac{2}{j^2} \int_0^j dj' \int_0^{j'} V_h(j'') dj'' \quad [4]$$

The absolute variogram of heterogeneity, $V_h(j)$, is equal to the relative variogram of the original variable in the time series, if the sample size is constant:

$$V_h(j) = \frac{1}{2(n-j)} \sum_{m=1}^{n-j} [h_{m+1} - h_m]^2, j = 1, 2, \dots, n/2 \quad [5]$$

Where j is the sample interval and h the heterogeneity. This heterogeneity h_m is given by:

$$h_m = \frac{(a_m - a_L)}{a_L} \frac{M_m}{M_{\bar{m}}} \quad m=1, 2, \dots, n \quad [6]$$

where a_m is the analytical result of the variable and a_L is the weighted mean of it in the time series. M_m is the size of the sample, $M_{\bar{m}}$ is the mean sample size and n is the total number of samples. In *samplinghelper*, once a scenario is selected, the program automatically generates a graph expressing $s^2(PSE)$ as a function of the number of samples, n . The user can define a realistic number of samples for the whole duration of the selected scenario, and the application stores the estimated errors.

2.1.2 Fundamental sampling error FSE

The volume of a sample is usually based on an analytical procedure, and fixed by regulations such as the ISO (ISO, 2006). In heterogeneous systems (e.g. suspended solids), the *FSE* is the variance between the concentrations in the unit particles of which the lot consists of. It is an intrinsic property of the material, which can be changed only if the particle size is changed. If the particle properties are known, in liquid samples the relative variance of the *FSE* can be estimated using the Poisson distribution:

$$s_{FSE}^2 = \frac{1}{n_p} \quad [7]$$

where n_p = number of particles in the sample.

If the size of analyte particles is d , its density ρ_p , mass concentration c , shape factor f , and size distribution factor g , then the relative variance of a sample volume V is

$$s_{FSE}^2 = \frac{f g \rho_p d^3}{c V} \quad [8]$$

Particle size is the size of the opening of a screen through which 95% of the material passes. The shape factor f is the ratio of the volume of the sampled particles having the characteristic dimension d to the volume of the cube having the same dimension. For spheroidal particles f is equivalent to 0.5, which is taken as the default value. The parameter g is the size distribution factor ($g = 0.25$ for large particle size distribution, and $g = 1$ for uniform particle sizes). Note that the heterogeneity will be higher (and so too the sample volume) if the concentrations of the substance of interest in the environment are low. If the compound is fully dissolved, the maximal particle size will be very small (few angstrom) and therefore the heterogeneity very small also. If the compound is adsorbed on particles, a particle size of 100 μm is a good proxy for environmental samples, using the particle as tracer.

In *Samplinghelper*, the user provides information about expected concentrations in the environment (average, minimum and maximum concentrations), and if the compound of interest is mainly dissolved or adsorbed on particles. Standard values are provided for the different factors. A plot that summarizes the s_{FSE}^2 equation is generated. The user can select a possible volume corresponding to his case (for example 500 ml), and the error generated is registered.

2.1.2 Grouping and segregation error GSE

The last equation for s_{FSE}^2 predicts smallest possible total sampling variance (variance of ideal sampling) only in two conditions: 1) a composite sample is collected picking single particles randomly from the lot; 2) the lot can be completely mixed before sampling. Segregation or grouping of analyte particles generates a grouping and segregation error, *GSE*. *GSE* is difficult to estimate and often in practice is assumed to be equal to the *FSE*.

2.1.3 Point materialization error PME

In the TOS, the *PME* is composed of three errors: the Increment Delimitation Error (*IDE*), the Increment Extraction Error (*IXE*) and the Increment and Sample Preparation Error (*ISE*), as defined by Gy (1996) and Minkinen and Esbensen (2009). To limit the *PME*, the sample must result from a correct selection (correctness condition), defined as: "A sampling is correct when it gives all elements, in the batch to be sampled, a uniform probability (> 0) of being selected". Samples extracted in sewers and rivers can be estimated as "correct" under the following conditions:

- High degree of turbulence = full mixing of the effluent at the sampling point (tested with dye tracer), or Reynolds number > 4000 (Graf and Altinakar, 1995);
- Peristaltic pumping device (no vacuum systems) (Haider and Haider, 1998), fluid speed in the tube $> 10 \times$ particle settling velocities;
- Varying sampling location between 0.6 – 0.8 times the water level;
- Short distance between sampling point and sampler (< 7 m)

Under these conditions, the standard estimated value for the materialization error is $\pm 20\%$ (De Heer, 1992; Rossi, 1998). If these conditions are relaxed, errors of 100% are possible, e.g., taking samples in big rivers after a conveyance before full mixing process or taking samples in a sewer pipe with laminar flow (small degree of turbulence). The *PME* is location and sampling system-specific and should therefore be estimated in situ. In the Web-application, the user can take the default value proposed, or provide an estimate.

2.1.4 Weighting error SWE

The concentration of a substance in the environment depends on the water flow. Errors in flow estimation directly influence the results. The volumetric error, expressed as "Sample Weighting error (*SWE*)" in the TOS, refers to errors in flow rate estimation. Two systematic sampling methods are commonly used: 1) Sampling increments at fixed time intervals and 2) sampling increments of the same volume when a fixed volume has passed the sampling point (volume-proportional sampling). If the increment sizes are not proportional to the flow-rate, and they are combined into a composite sample, the first method gives a composite sample that is likely to be biased by the weighting error *SWE*. A systematic error component that cannot be reduced by increasing the number of increments can be generated (Minkinen, 2007). The second method, volume-proportional sampling, gives an unbiased mean both for correlated and uncorrelated data sets.

The volumetric error can be estimated in different ways at the measuring site. A volumetric mass balance can help to estimate the reliability of the measurements. As an approximation, we propose 20% as standard value for the volumetric error in an ideal case (volume-proportional sampling, multiple checks of the water flow measuring stations, establishment of a volumetric mass balance, comparison with other upstream or downstream stations, etc.). Note that the errors of measurement are not similar to the (usually very low) instrument errors. Local conditions (turbulence) and rating curves are sources of much higher uncertainties. As for the *PME*, in *Samplinghelper*, user-defined estimates of *SWE* can be entered.

2.1.5 Total analytical error TAE

The *TAE* is the total of the errors made during the analytical process in the laboratory. It includes errors due to degradation processes during the storage of the samples. Analytical errors depend on numerous factors, like kind of matrix (lake water, river water, wastewater...), expected concentrations (usually bigger errors with lower concentrations), analytical methods, human factors, etc. Laboratories

are usually certified and deliver results within a given precision range. Analytical errors are expected to be much smaller than sampling errors. In our application, instead of using the defined laboratory precision (usually indicated in the results protocol), we prefer to use the results of ring tests or inter-laboratory calibration to evaluate the analytical errors (ISO 13528, 2005). In ring tests, all laboratories receive the same sample. The results of one laboratory could be “precise and reproducible”, but not reliable. In our application, we provide results of ring tests for different compounds (Table 1). The user can select one of these values or estimate the error in another way, using for example dedicated software (UncertaintyManager ©, Ellison et al, 2002).

Compound	Media	Target value	Uncertainty range	References
Ammonia*	Lake water	0.028 mg N/l	22%	Strawczynski, 2008
N _{tot} *	Lake water	1.69 mg N/l	17%	
P-PO4 *	Lake water	0.046 mg P/l	8%	
P _{tot} *	Lake water	0.120 mg P/l	6%	
Ibuprofen (pharmaceutical)	Wastewater	1867 ng/l	147%	Farré et al. 2008
Ibuprofen* (pharmaceutical)	River water	675 ng/l	57%	
Ibuprofen* (pharmaceutical)	MiliQ water	225 ng/l	27%	

Table 1: Example of inter-laboratory results for some compounds and different media (Lake water, wastewater, river water and MiliQ water) *fortification of the samples with known concentrations.

2.1.6 Global estimation error GEE

The *GEE* summarizes the results of the foregoing errors. The errors are assumed to be independent, and so the *GEE* is RMS of the individual errors. In the application, a graphic illustrates the relative contribution of each step to the total error, helping to convey where uncertainties arise in the sampling plan. A practical outcome is that investment can then be appropriately directed to improve the results.

3 CASE STUDY

The results from a six-week measuring campaign of flow rate and ammonia at the Ergolz WWTP (Canton Baselland, Switzerland) are presented to illustrate the software. For flow measurements, an ultrasonic probe (Milltronics PL-372) was installed near the WWTP entrance, upstream of a combined sewer overflow (CSO), with a sampling rate of 1/minute. The relation between water level and flow was established with a Venturi channel installed downstream of the CSO. During rain events, the amount of water discharged to the CSO tank was recorded in order to be able to calibrate it also at high water flows.

Ammonia was measured each minute with an ion-selective membrane device (AmmonLyt 700 IQ ammonium sensor). The probe was mounted on a floating device at the WWTP inlet, allowing for measurement during turbulent flow. Laboratory ammonia analyses were performed on manual samples of one litre five to ten minutes after sampling (Dr. Lange #LCK303). The same sample was analyzed 24 hours later, after complete decomposition of any urea present. On the basis of these manual measurements, the on-line ammonia measurements were converted to total urea and ammonia N (linear correlation, $R^2 = 0.92$). Since flow and concentration were both measured every minute, the resulting load of ammonia/urea-N was obtained with the same high resolution. Data collected are illustrated in Figure 2. They are automatically generated on the webpage to check visually the information before going to the next step.

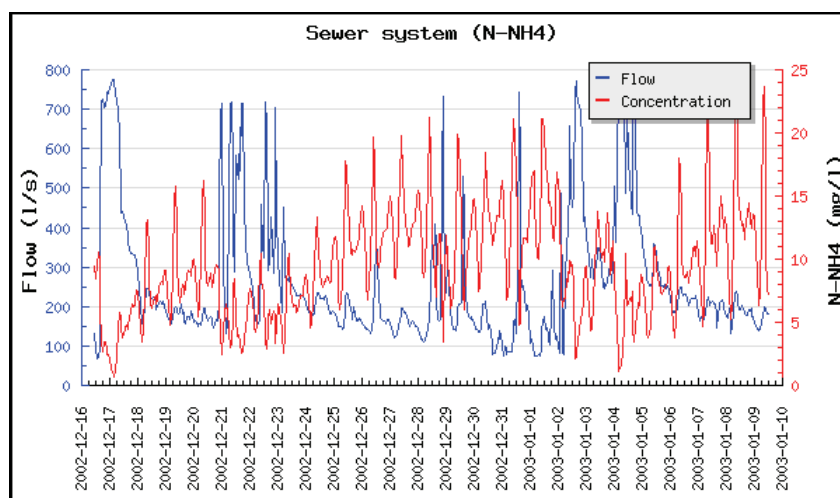


Figure 2 : Scenario implemented in the Web-application: Concentrations of N-NH₄⁺ and flow rate measured continuously at the entry of the WWTP Ergolz1 during three weeks

A variogram is calculated from the measurements and displayed to estimate the *PSE*. The user can determine the type of sampling strategy and then define the number of samples planned. The application displays the estimated error for the selected number of samples.

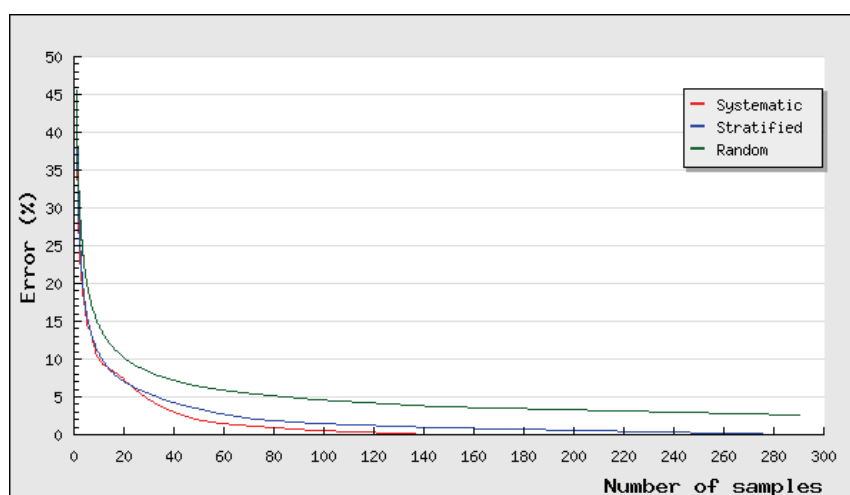


Figure 3: Point selection error for the scenario selected. In this example, 24 samples would be taken in the same period, in a systematic way, generating an error of 6 %.

In the next step, users provide information for estimating *FE*. Information about environmental concentrations is required. Usually, such information may be found in the literature. One important point is whether the target compound is adsorbed on particles or completely dissolved. Ammonia is mainly dissolved, so a very small value for the parameter "*d*" is selected. Results change markedly for phosphorous or heavy metals, as these compounds are mainly adsorbed on particles. Standard values are proposed for other granulometric factors. The results are automatically generated, illustrating the standard deviation of the fundamental error as it varies with the mass in millilitre (i.e., gram) of the sample. An uncertainty range, accounting for variations concentrations and uncertainties is also presented. For this example, the error is very small for dissolved compound like ammonia (< 1%), as the heterogeneity is in that case very small also.

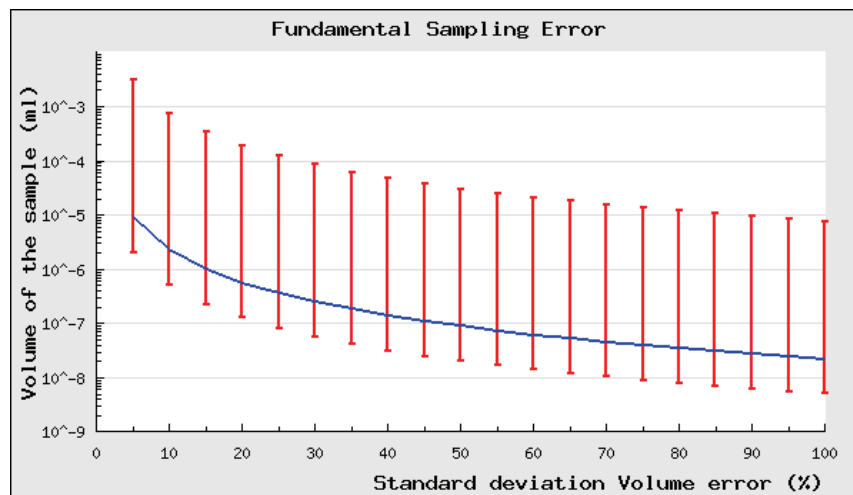


Figure 4: Fundamental sampling error for the scenario selected. In this example, 200-ml samples are regularly taken, generating negligible uncertainty as ammonia is a fully dissolved compound.

In this example, we optimized the sampling location in order to minimize the *PME*. The use of a floating device permits optimal positioning of the measuring point, at 0.8 times the water level in the middle of the cross-section. The degree of turbulence was sufficient to ensure a good flow mixing, especially during rain events. As measurements were made on-line, the sampling extraction error and the sample preparation error were neglected. We can estimate the increment delimitation error as very small, a value of 5% was retained for this example.

The weighting error was estimated by comparing our measurement with measurements conducted at the WWTP in the Venturi channel. The high degree of turbulence in the flow generated variations in the water level measurements of ± 1 cm during dry weather to ± 5 cm during rain events. Taking into account also uncertainties in the rating curve, a global estimation of 15% for the *SWE* was estimated.

For the total analytical error, results of intercalibration studies and probe calibration are considered. For ammonia, inter-calibration results for ammonia revealed a precision of 22%. For the on-line probe, we have also to consider the differences between the signal measured by the probe and chemical analysis conducted in the laboratory. The explained variance between the two measurements was estimated at $R^2 = 0.92$. The total analytical error was estimated to be 23%.

Finally, the last page of the program summarizes the different results, considering all uncertainty contributions in an independent manner (Figure 5). We can see, for this example, that the main source of error is linked with the on-line probe measuring device and flow measurements.

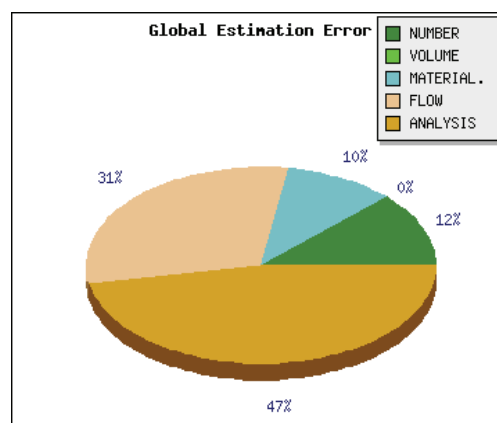


Figure 5: Illustration of the relative contribution of all sampling steps to the global estimation error (*GEE*), 29% in this case. The biggest contribution is related to the analytical step and flow measurements for ammonia measurements.

4 CONCLUSIONS AND PERSPECTIVES

Sampling represents a key issue for obtaining accurate and precise information. In this paper, an innovative tool is presented to help practitioners to better understand and describe different kind of errors that can be generated during sampling for water monitoring. Unfortunately it still happens too often that results of sampling campaigns are not usable due to an inappropriate sampling plan, a situation that this software helps avoid. In the illustrated example, on-line measurements with new technical devices placed in optimal condition at the inlet of a WWTP generated an uncertainty evaluated at 29%. Good mixing conditions were present for ammonia, generating uncertainties in the flow measurements. High measurement frequency was possible with the on-line system, but the analytical precision was relatively poor in comparison with classical analytical techniques.

More scenarios are needed to cover "standard" pattern of water pollution, accounting for different flow conditions (small/medium/big rivers or small/medium/big sewer networks) and different pollution patterns (dissolved/particular compounds, punctual/diffuse/accidental source of pollution, etc.). The goal of this application is not to cover all these possible situations, but to alert users to a "What if..." situation. Local conditions can be completely different from the proposed scenarios. For that reason, sampling design must be considered as an iteration: a pilot-campaign of measurement will help to better estimate the potential pattern, and the sampling strategy can be optimized accordingly.

Efforts should also be applied to instrumentation for liquid sampling and flow measurements. Commercially available sampling devices have some limitations. For example, an iso-kinetic water sampling device does not exist to our knowledge. An "ideal" sampling system for water monitoring still needs to be developed. If the same investment is made in sampling devices as for analytical techniques, such systems would be certainly soon on the market.

The application of the TOS on a Web application aims to encourage practitioners, regulatory bodies and scientists to plan robust sampling campaigns. It is also a good learning tool for students.

LIST OF REFERENCES

- Bertrand-Krajewski J.L., Bardin J.P., Mourad M. and Béranger Y. (2003). Accounting for sensor calibration, data validation, measurement and sampling uncertainties in monitoring urban drainage systems. *Water Science and Technology* 47(2), 95–102.
- Chèvre N, Singer H, Müller S, Müller E. 2004. Evaluation du risque des pesticides dans les eaux courantes en Suisse (Risk assessment of pesticides in Swiss receiving waters). *Gaz Wasser Abwasser* 10, 739-751.
- Christakos G. and Olea R.A. (1992). Sampling design for spatially distributed hydrogeologic and environmental processes. *Advances in Water Resources* 15, 219-237.
- De Heer J. (1992). Etude de l'échantillonnage systématique et proportionnel au débit par temps sec des eaux usées dans des égouts non visitables. PhD thesis n°1078, EPFL, Lausanne, Switzerland.
- Ellison S.L.R., King B., Rösslein M., Salit M., Williams A. (2002). Traceability in Chemical Measurement. Eurachem/CITAC Draft Guide.
- Farré M., Petrovic M., Gros M., Kosjek T., Martinez E., Heath E., Osvald P., Loos R., Le Menach K., Budzinskif H., De Alencastro F., Müller J., Knepper T., Fink G., Ternes T.A., Zuccato E., Kormali P., Gans O., Rodil R., Quintana J.B., Pastori F., Gentili A., Barcelo D. (2008). First interlaboratory exercise on non-steroidal anti-inflammatory drugs analysis in environmental samples. *Talanta* 76, 580–590.
- Graf W.H., Altinakar M.S. (1995). Hydrodynamique : une introduction. *Traité de génie civil* 14, PPUR (ed), EPFL, Lausanne, Switzerland.
- Gy P. (1988). Hétérogénéité, échantillonnage, homogénéisation. Ensemble cohérent de théories. Masson, Paris.
- Gy P. (1992). Sampling of Heterogeneous and Dynamic Material Systems - Theory of Heterogeneity, Sampling and Homogenizing. Elsevier, Amsterdam.
- Gy P. (1996). L'échantillonnage des lots de matières en vue de leur analyse. In: Collection mesures physiques, science de l'ingénieur. Masson, Paris.
- Haider R. and Haider S. (1998). Overestimated COD values for raw wastewaters extracted by vacuum sampling devices due to sedimentation effects. *Wat. Sci. Tech.* 37(12), 293-300.
- Isidoro D., Quilez D., Aragües R. (2003). Sampling strategies for the estimation of salt and nitrate loads in irrigation return flows: La Violada Gully (Spain) as a case study. *Journal of Hydrology* 271, 39–51.

- ISO 13528 (2005). Statistical methods for use in proficiency testing by interlaboratory comparisons. Report, First edition 2005-09-01. ISO, Geneva, Switzerland.
- Johnes P.J. (2007). Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. *Journal of Hydrology* 332, 241-258.
- Littlewood I.G. (1995). Hydrological regimes, sampling strategies and assessment of errors in mass load estimates for United Kingdom rivers. *Environment International* 21(2), 211-220.
- Minkinen P. (2004). Practical applications of sampling theory. *Chemometrics and Intelligent Laboratory Systems* 74, 85-94.
- Minkinen P. (2007). Weighting error – Is it significant in process analysis? In João Felipe Costa and Jair Koppe, eds., WCBS3, Third World Conference on Sampling and Blending, Porto Alegre, Brazil.
- Minkinen P., Esbensen K.H. (2009). Grab vs. composite sampling of particulate materials with significant spatial heterogeneity—A simulation study of “correct sampling errors”. *Analytica Chimica Acta* 653, 59-70.
- Ort C. and Gujer W. (2006). Sampling for representative micropollutant loads in sewer systems. *Water Science & Technology* 54(6-7), 169–176.
- Paakkunainen M., Reinikainen S.P., Minkinen P. (2007). Estimation of the variance of sampling of process analytical and environmental emissions measurements. *Chemometrics and Intelligent Laboratory Systems* 88, 26–34.
- Petersen L., Minkinen P., Esbensen K.H. (2005). Representative sampling for reliable data analysis: Theory of Sampling. *Chemometrics and Intelligent Laboratory Systems* 77, 261- 277.
- Pitard F. F. (1993). *Pierre Gy's Sampling Theory and Sampling Practice*. 2nd ed., CRC Press.
- Pitard F.F. (2009). *Pierre Gy's Theory of Sampling and C.O. Ingamell's Poisson Process Approach – Pathways to Representative Sampling and Appropriate Industrial Standards*. PhD Thesis, Aalborg University, Denmark.
- Rossi L (1998). *Qualité des eaux de ruissellement urbaines*. PhD thesis n°1789, EPFL, Lausanne, Switzerland.
- Strawczynski A. (2008). *Analyses comparatives inter-laboratoires (Comparative interlaboratory analysis)*. Rapp. Comm. Int. Prot. Eaux Léman contre pollut., Campagne 2007, 163-175.